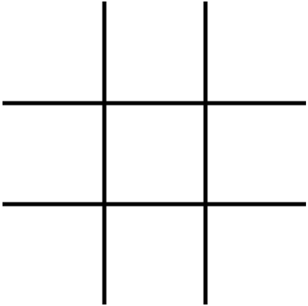
Why Learning?

- Usually: Equilibrium concepts
- Most prominent: Nash-Equilibrium and refinements
- but...

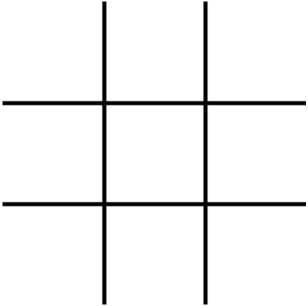
Why Learning?

Lets play a game of Tic Tac Toe, you start

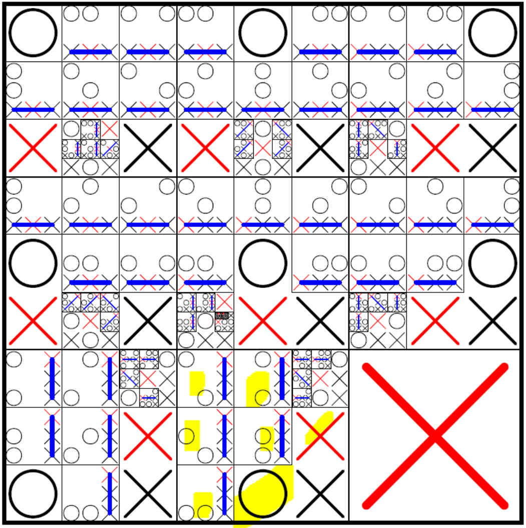


Why Learning?

Lets play a game of Tic Tac Toe, you start



Did you calculate the (subgame perfect) Nash-EQ?



Why Learning?

- Finding the equilibrium already quite complicated in a very simple game
- Even harder (practically impossible!) in other games: checkers, chess, etc ...
- Nash-EQ assumes that all players are able to calculate the EQ perfectly

  Some learning theories involve considerably less calculations (but not all)
- Even if you can calculate the EQ, are you sure your opponent can as well?

Why Learning?

Imagine that during the last 2 years, you have gone to the opera on each first monday in the month with your partner. It is the first monday of this month. As usual, there is an opera being conducted and a football match being played in your town. You did not talk to your partner about the plans for today (and your cellphone has died). The following matrix describes the payoffs of you and your partner.

	Opera	Football
Opera	3,2	0,0
Football	0,0	2,3

Battle of the Sexes 1

- Would you go to the opera or the football match?

Why Learning?

Imagine that during the last 2 years, you have gone to the opera on each first monday in the month with your partner. It is the first monday of this month. As usual, there is an opera being conducted and a football match being played in your town. You did not talk to your partner about the plans for today (and your cellphone has died). The following matrix describes the payoffs of you and your partner.

	Opera	Football
Opera	3,2	0,0
Football	0,0	2,3

Battle of the Sexes 1

- Would you go to the opera or the football match?
- Both going to the match and going to the opera are part of a Nash-EQ, but one seems intuitively more likely

Why Learning?

- Reasons not to use Nash-EQ
  - Strong assumptions: Full rationality, common knowledge
  - Prediction not always correct (other theories have to prove they are better!)
  - Nash-EQ is only “as-if”, what is really happening in the brain?
- Learning theories abandon the Nash-EQ assumptions, similar the non-EQ concepts we heard about last week, but also incorporate history
  - Learning theories only make sense in the environment of a repeated interaction
  - Learning across games??

Information used

Learning theories use information about past events to come up with a prediction/description of present play.
What information is used?

- Last rounds strategy of other player(s) + own payoff matrix
  => Cournot Adjustment

Information used

Learning theories use information about past events to come up with a prediction/description of present play.
What information is used?

- Last rounds strategy of other player(s) + own payoff matrix => Cournot Adjustment
- Past strategies of other player(s) + own payoff matrix => Fictitious Play

Information used

Learning theories use information about past events to come up with a prediction/description of present play.
What information is used?

- Last rounds strategy of other player(s) + own payoff matrix => Cournot Adjustment
- Past strategies of other player(s) + own payoff matrix => Fictitious Play
- Past own strategies and associated payoffs => Reinforcement Learning

Information used

Learning theories use information about past events to come up with a prediction/description of present play.
What information is used?

- Last rounds strategy of other player(s) + own payoff matrix => Cournot Adjustment
- Past strategies of other player(s) + own payoff matrix => Fictitious Play
- Past own strategies and associated payoffs => Reinforcement Learning
- Past own and others strategies + own payoff matrix => Experienced Weighted Attractions learning

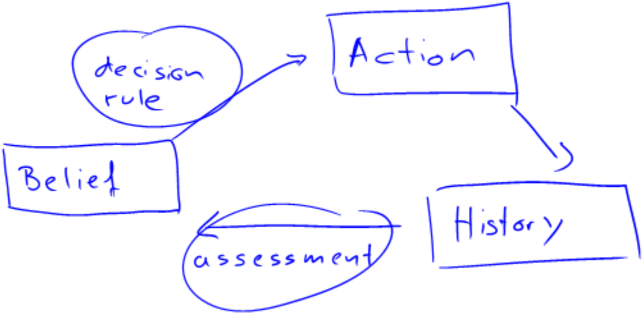
Information used

Learning theories use information about past events to come up with a prediction/description of present play.
What information is used?

- Last rounds strategy of other player(s) + own payoff matrix => Cournot Adjustment
- Past strategies of other player(s) + own payoff matrix => Fictitious Play
- Past own strategies and associated payoffs => Reinforcement Learning
- Past own and others strategies + own payoff matrix => Experienced Weighted Attractions learning
- Last rounds payoffs and strategies => Imitation, (Replicator Dynamic)

II. 1. Belief-based learning

1.1 Cournot best reply process
1.2 Fictitious play



Cournot adjustment

A very simple learning model: Cournot adjustment/Myopic best responses (book: Fudenberg/Levine, The Theory of Learning in Games, chapter 1)

- Based on the known notion of **best response**
- Each player plays a best response to the other player's last round strategy
- Adjustment is simultaneous (both adjust at the same time) or sequential (always one player is "locked-in" for one round)
- Example: Cournot duopoly

Cournot adjustment

Notation

- The players are two firms, who both set quantities and want to maximise their payoff, strategies $s^i \in S^i$, utility $u^i(s_i, s_{-i})$
- Firm i's best response function is: $BR^i(s^{-i}) = argmax_{s^i} u^i(s^i, s^{-i})$
- When both firms simultaneously update their quantity, then the quantity of a firm is given by: $s^i_t = BR^i(s^{-i}_{t-1})$

Cournot adjustment

- Why is Cournot adjustment a learning process and not an equilibrium concept?
- Path dependence: History matters
  - Perhaps a drawback: Needs some starting values
  - Different starting values can lead to different end result
- Rationality of the other player assumed to be very weak:
  - Static behavior
  - What is problematic about this assumption?

Cournot adjustment

- What information is needed to calculate the new strategy?

Cournot adjustment

- What information is needed to calculate the new strategy?
  - Last round opponent strategy + own payoff function
    => Only one round of strategies important for future play

Cournot adjustment

- What information is needed to calculate the new strategy?
  - Last round opponent strategy + own payoff function
    => Only one round of strategies important for future play
- Where does the adjustment process converge to?
- What happens at the Nash-EQ?
- Does it always converge to the Nash-EQ?

Stability

Take a system of were players are repeatedly making decisions to play one of their strategies according to some (fixed) rule based on the history of the game. E.g. The cournot adjustment process.

- *state:* All variables needed to determine which strategies are played, $\theta_t \in \Theta$

Stability

Take a system of were players are repeatedly making decisions to play one of their strategies according to some (fixed) rule based on the history of the game. E.g. The cournot adjustment process.

- *state:* All variables needed to determine which strategies are played, $\theta_t \in \Theta$ Last round's quantities

Stability

Take a system of were players are repeatedly making decisions to play one of their strategies according to some (fixed) rule based on the history of the game. E.g. The cournot adjustment process.

- *state:* All variables needed to determine which strategies are played, $\theta_t \in \Theta$ Last round's quantities
- law of motion: $\theta_{t+1} = f_t(\theta_t)$

Stability

Take a system of were players are repeatedly making decisions to play one of their strategies according to some (fixed) rule based on the history of the game. E.g. The cournot adjustment process.

- *state:* All variables needed to determine which strategies are played, $\theta_t \in \Theta$ Last round's quantities
- law of motion: $\theta_{t+1} = f_t(\theta_t)$ Best Response to other player's quantity

Stability

Take a system of were players are repeatedly making decisions to play one of their strategies according to some (fixed) rule based on the history of the game. E.g. The cournot adjustment process.

- *state:* All variables needed to determine which strategies are played, $\theta_t \in \Theta$ Last round's quantities
- law of motion: $\theta_{t+1} = f_t(\theta_t)$ Best Response to other player's quantity
- Value at time *t*, when the initial state was θ_0: $F_t(\theta_0)$

Stability

Take a system of were players are repeatedly making decisions to play one of their strategies according to some (fixed) rule based on the history of the game. E.g. The cournot adjustment process.

- *state:* All variables needed to determine which strategies are played, $\theta_t \in \Theta$ Last round's quantities
- law of motion: $\theta_{t+1} = f_t(\theta_t)$ Best Response to other player's quantity
- Value at time t, when the initial state was θ_0: $F_t(\theta_0)$
- *Steady state (absorbing)*: state does not change any more, $F_t(\hat{\theta}) = \hat{\theta}, t > 0$

Stability

Take a system of were players are repeatedly making decisions to play one of their strategies according to some (fixed) rule based on the history of the game. E.g. The cournot adjustment process.

- *state:* All variables needed to determine which strategies are played, $\theta_t \in \Theta$ Last round's quantities
- law of motion: $\theta_{t+1} = f_t(\theta_t)$ Best Response to other player's quantity
- Value at time t, when the initial state was θ_0: $F_t(\theta_0)$
- *Steady state (absorbing)*: state does not change any more, $F_t(\hat{\theta}) = \hat{\theta}, t > 0$
- (Strict) Nash-EQs are steady states of the Cournot adjustment process

Stability

Do we reach the steady state?

- *(Lyapunov) stable steady state:* $\forall U$ of $\hat{\theta}$, $\exists U_1$ of $\hat{\theta}$, such that if $\theta_0 \in U_1$, then $F_t(\theta_0) \in U, t > 0$
  - U, U_1 Neighborhoods of $\hat{\theta}$

Stability

Do we reach the steady state?

- *(Lyapunov) stable steady state:* $\forall U$ of $\hat{\theta}$, $\exists U_1$ of $\hat{\theta}$, such that if $\theta_0 \in U_1$, then $F_t(\theta_0) \in U, t > 0$
    - U, U_1 Neighborhoods of $\hat{\theta}$
- *Asymptotically stable steady state*: In addition to above, if $\theta_0 \in U_1$, then $lim_{t \to \infty} F_t(\theta_0) = \hat{\theta}$

Stability

Do we reach the steady state?

- *(Lyapunov) stable steady state:* $\forall U$ of $\hat{\theta}$, $\exists U_1$ of $\hat{\theta}$, such that if $\theta_0 \in U_1$, then $F_t(\theta_0) \in U, t > 0$
  - U, U_1 Neighborhoods of $\hat{\theta}$
- *Asymptotically stable steady state*: In addition to above, if $\theta_0 \in U_1$, then $lim_{t\to\infty} F_t(\theta_0) = \hat{\theta}$
- *Basin of attraction*: Set of all points θ_0 such that $lim_{t\to\infty} F_t(\theta_0) = \hat{\theta}$

Stability

Do we reach the steady state?

- *(Lyapunov) stable steady state:* $\forall U$ of $\hat{\theta}$, $\exists U_1$ of $\hat{\theta}$, such that if $\theta_0 \in U_1$, then $F_t(\theta_0) \in U, t > 0$
  - U, U_1 Neighborhoods of $\hat{\theta}$
- *Asymptotically stable steady state*: In addition to above, if $\theta_0 \in U_1$, then $lim_{t\rightarrow\infty} F_t(\theta_0) = \hat{\theta}$
- *Basin of attraction*: Set of all points θ_0 such that $lim_{t\rightarrow\infty} F_t(\theta_0) = \hat{\theta}$
- In Cournot example: Nash-EQ asymptotically stable steady state, basin of attraction equal to entire state space => *globally stable*

Fictitious Play

- General idea: Best response to all previous rounds, not just last round
- Still path dependent
- Opponent still assumed to be static
  - But now any distribution, not just one strategy
- Book: Fudenberg/Levine, The Theory of Learning in Games, chapter 2

History of Fictitious Play

- Developed by Brown, 1951
- Originally not developed as a learning theory:
  "*It is the purpose of this chapter to describe (...) a simple iterative method for approximating to solutions of discrete zero-sum games*"
- To be useful as approximisation, should converge to EQ
- Long search for classes of games where Ficititous Play converges
- But: Does not always converge
- Later re-interpreted from EQ approximisation device to learning theory (that is, non-EQ behavior)

2 Player Fictitious Play

- 2 players, repeatedly playing the same game
- Finite strategy spaces S_1, S_2
- Players form *assessments* over the distribution of opponents actions
  - Imagine the game was already played some rounds and players kept a statistic of which actions the opponent played
  - As if players believe that the opponent is always playing according to a random draw from a fixed(!) distribution
  - Need some initial weights to deterime play in first round (prior beliefs)
  - Fictitious Play a form of *Belief-based learning*

2 Player Fictitious Play

- Initial weight function: $\kappa_0^i : S^{-i} \to \mathbb{R}_+$
- Each time the opponent plays a strategy s_{-i}, 1 is added to the weight of that strategy
- $\kappa_t^i(s^{-i}) = \kappa_{t-1}^i(s^{-i}) +$
  - 1 if $s_{t-1}^{-i} = s^{-i}$
  - 0 if $s_{t-1}^{-i} \neq s^{-i}$
- κ counts empirical distribution of actions used by other player (plus initial weight)

2 Player Fictitious Play

- Players act as if empirical distribution perfectly predicts future play of the opponent
- Probability assigned to player $-i$ playing s^{-i} at date t: $\gamma_t^i(s^{-i}) = \frac{\kappa_t^i(s^{-i})}{\sum_{\tilde{s}^{-i} \in S^{-i}} \kappa_t^i(\tilde{s}^{-i})}$
  - Probabilty equal to relative weight
- Fictitious Play demands that each player plays a best response to his beliefs: $BR^i(\gamma_t^i)$
  - Not unique: There could be more than one best response
  - Chose a pure strategy is indifferent
  - Indifference will not happen with generic payoffs

2 Player Fictitious Play: Example

We look again at the matching pennies game from the last lecure:

	Left (48)	*Right* (52)
Top (**48**)	80, 40	40, 80
Bottom (**52**)	40, 80	80, 40

- Assume the initial weights are
  - Player 1 (row): 1.5 on left and 1 on right
  - Player 2 (column): 1.5 on top and 1 on bottom
- Calculate the weights for each player for the first 5 rounds of play and calculate which action the players will chose in each round

2 Player Fictitious Play: Convergence

- Strict Nash-EQ are steady states of Fictitious Play
- Any pure-strategy steady state of Ficititous Play must be a Nash-EQ
- If the empirical distribution of players choices converges, the corresponding strategy profile must be a Nash-EQ

2 Player Fictitious Play: Convergence

When does Fictitious Play converge?
If the game is generic and:

- 2x2, Robinson 1951, extended to 2xn by Berger, 2005
- Zero-sum, Miyasawa 1951
- Is solvable by iterated strict dominance, Nachbar 1990
- Strategic complements + decreasing marginal return, Krishna 1992
- 3x3 + stragic complements, Hahn, 1999
- and more

However ...

2 Player Fictitious Play: Cycles

Example by Fudenberg/Krebs

	A	B
A	0,0	1,1
B	1,1	0,0

- What would happen under Cournot adjustment?
- Under Fictitious Play?

2 Player Fictitious Play: Cycles

Example by Fudenberg/Krebs

	A	B
A	0,0	1,1
B	1,1	0,0

- What would happen under Cournot adjustment?
- Under Fictitious Play?
- =>Cycle!
- FP closely related to the Cournot best response process

2 Player Fictitious Play: Cycles

Why are cycles problematic?

- Can stop convergance to Nash-EQ
- Not plausible, given the interpretation that opponent draws from a distribution
- Can lead to payoffs that are “off” from what we would expect, even if strategy choices approach Nash-EQ

Fictitious Play: Extensions

- Sequential instead of Simulaneous adjustment (Original Brown paper)
- FP in continuous time: Often results similar, sometimes mathematically easier
- More than 2 players: Have to decide how to track opponents play, separately or joint
- Specific assumptions about tie-breaking (non-generic games)
- Stochastic element even without ties
- weighted FP: Puts more weight on recent history
- EWA (later lecture)

Fictitious Play: Experimental Evidence

Huck/Normann/Oechssler (2002)

- 4 firm Cournot oligopoly
- Test predictive power of BR (Cournot adjustment), FP (Fictitious Play) and AV (Imitate the average)
- $z = \frac{quantity_t - quantity_{t-1}}{prediction - quantity_{t-1}}$
  - $z = 1$ perfect prediction
  - $z < 0$ even direction of adjustment wrong
- Data available

Fictitious Play: Experimental Evidence

Table 3. Hit ratios

Treatment		$z < 0$	$0 \leq z < .8$	$0.8 \leq z < 1$	$z = 1$	$1 < z \leq 1.2$	$z > 1.2$
A	BR	21.5	52.0	4.6	5.2	4.1	12.6
	FP	49.7	22.5	4.2	0	4.7	18.8
	AV	35.3	46.3	2.5	0.5	2.4	13.0
B	BR	26.4	43.0	5.4	5.4	3.1	16.6
	FP	46.0	22.1	5.3	0	4.3	22.3
	AV	35.7	44.6	4.1	1.0	2.3	12.3

Fictitious Play: Experimental Evidence

Cheung/Friedman (1997)

- Use several games (Hawk-Dove, Coordination, Battle of Sexes, Buyer-Seller)
- Test 3 parameter model, which includes Cournot adjustment and Fictitious Play as special cases

Type	Fickle	Cournot	Adaptive	Fictitious	Imprint	Uninformative
BoS	4	51	6	18	0	18
B–S	4	37	9	17	0	25
Co	5	25	6	18	1	22
HD	15	31	8	31	1	41
Total	28	144	29	84	2	106

Fictitious Play: Experimental Evidence

- Cournot somewhat better than Fictitious Play
- But: Undescribed players?
- Other (non-belief based) learning theories?

Fictitious Play: Experimental Evidence

- Cournot somewhat better than Fictitious Play
- But: Undescribed players?
- Other (non-belief based) learning theories?

Duersch/Kolb/Schipper/Oechssler (2010)

- Cournot Duopoly vs computer player
- MSD: Mean Squared Deviation, MSD$=0$ means perfect prediction

Fictitious Play: Experimental Evidence